1.0

4.5
5.0
5.6

2.8

2.5

3.2

2.2

3.6

1.1

4.0

2.0

1.8

1.25

1.4

1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A117439

DTIC FILE COPY

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AIM 678	2. GOVT ACCESSION NO. AD-A117439	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Learning By Augmenting Rules and Accumulating Censors		5. TYPE OF REPORT & PERIOD COVERED MEMORANDUM
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Patrick H. Winston		8. CONTRACT OR GRANT NUMBER(s) N00014-80-C-0505
9. PERFORMING ORGANIZATION NAME AND ADDRESS Artificial Intelligence Laboratory 545 Technology Square Cambridge, Massachusetts 02139		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Research Projects Agency 1400 Wilson Blvd Arlington, Virginia 22209		12. REPORT DATE July 1982
		13. NUMBER OF PAGES Pages, 22
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Office of Naval Research Information Systems Arlington, Virginia 22217		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is unlimited.

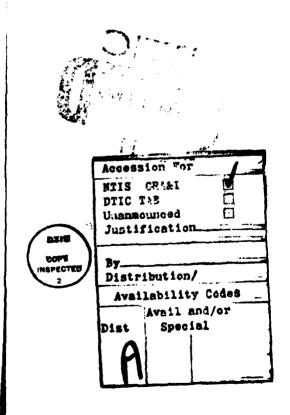17. DISTRIBUTION STATEMENT (of the abstract entered in Block 30, if different from Report)

DTIC
ELECTE
JUL 23 1982
H

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

LEARNING, ARTIFICIAL INTELLIGENCE, ANALOGY

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper is a synthesis of several sets of ideas: ideas about learning from precedents and exercises, ideas about learning using near misses, ideas about generalizing if then rules, and ideas about using censors to prevent procedure misapplication. The synthesis enables two extensions to an implemented system that solves problems involving precedents and exercises and that generates if-then rules as a byproduct. These extensions are as follows: If then rules are augmented by if plausible conditions creating augmented if then rules. An augmented if then rule is blocked whenever →

88 07 20 000

facts in hand directly deny the truth of an if plausible condition. When an augmented if then rule is used to deny the truth of an if plausible condition, the rule is called a censor. Like ordinary augmented if then rules, censors can be learned.

Definition rules are introduced that facilitate graceful refinement. The definition rules are also augmented if-then rules. They work by virtue of if-plausible entries that capture certain nuances of meaning different from those expressible by necessary conditions. Like ordinary augmented if-then rules, definition rules can be learned. The strength of the ideas is illustrated by way of representative experiments. All of these experiments have been performed with an implemented system.

ARTIFICIAL INTELLIGENCE LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

AIM 678   May 1982

LEARNING BY AUGMENTING RULES AND ACCUMULATING CENSORS

by

Patrick H. Winston

Abstract

This paper is a synthesis of several sets of ideas: ideas about learning from precedents and exercises, ideas about learning using near misses, ideas about generalizing if-then rules, and ideas about using censors to prevent procedure misapplication.

The synthesis enables two extensions to an implemented system that solves problems involving precedents and exercises and that generates if-then rules as a byproduct. These extensions are as follows:

- If-then rules are augmented by *if-plausible* conditions, creating *augmented if-then rules*. An augmented if-then rule is blocked whenever facts in hand directly deny the truth of an if-plausible condition. When an augmented if-then rule is used to deny the truth of an if-plausible condition, the rule is called a *censor*. Like ordinary augmented if-then rules, censors can be learned.

- Definition rules are introduced that facilitate graceful refinement. The definition rules are also augmented if-then rules. They work by virtue of *if-plausible* entries that capture certain nuances of meaning different from those expressible by necessary conditions. Like ordinary augmented if-then rules, definition rules can be learned.

The strength of the ideas is illustrated by way of representative experiments. All of these experiments have been performed with an implemented system.

—\—

KEY IDEAS

This paper builds primarily on a previous paper that introduced a theory of learning from precedents and exercises using *constraint transfer* [Winston 1981]. The theory addresses the analogy process at work when we exploit past experience in fields like Management, Political Science, Economics, Medicine, and Law, as well as from everyday life.

Two extensions to the theory are described. Work on the first extension was stimulated by some of the apparent blunders of the extant system. Work on the second extension was stimulated by some problems encountered in making definitions.

After a brief review of the overall theory, I present an example showing that the rules generated by the unextended learning system can be misapplied. Next, I discuss various solutions to the misapplication problem, including the introduction of censors. At this point *augmented if-then rules* are discussed. Each augmented if-then rule contains not only *if* and *then* parts, but also an *if-plausible* part. Before a rule acts, censors determine if any existing facts directly demonstrate that an *if-plausible* relation is false. If so, the rule is *blocked.*

This leads to the development of definition rules based on augmented if-then rules and a discussion of their relevance to the problem of concise definition versus unlimited nuance.

Next, it is shown that censors can block censors and that censors can be learned, both by precedent and exercise and by near miss.

Finally, I describe precedents for this work itself, including ideas that stimulated what I have done, such as Minsky's ideas on the role of censors in problem solving [Minsky 1980], as well as other ideas that I reinvented or borrowed from as my work progressed, such as Goldstein and Grimson's ideas on generalizing if-then rules [1977].

There are references throughout to an implemented system that actually does acquire and use censors. This implemented system inherits some key ingredients from previous work:

o   **Analogy-based reasoning using** *constraint transfer.* Analogy requires the ability to determine how two situations that are similar in some respects may be similar in other respects as well. Here the determination is done by transferring constraints from the precedent situation to the exercise situation.

o   **Learned if-then rules.** In contrast to current practice in Knowledge Engineering, if-then rules emerge automatically as problems are solved. Teachers supply precedents and exercises, leaving the work of formulating the if-then rules to the system.

o   Actor-object representation. Situations are represented using relations between pairs
    of situation parts. Supplementary descriptions can be attached to the relations when
    elaboration is needed.

o   Importance-dominated matching. The similarity between two situations is measured
    by finding the best possible match according to what is important in the situations.
    Importance is determined by causal connections in the situations themselves. Causal
    connection is viewed as a common importance-determining constraint.

WHAT IS TO BE UNDERSTOOD

Let us begin by reviewing the sort of task performed by the theory as previously reported.
Consider the following precis of *Macbeth*, given by a teacher as a precedent:

> MA is a story about Macbeth, Lady-macbeth, Duncan, and Macduff. Macbeth
> is an evil noble. Lady-macbeth is a greedy, ambitious woman. Duncan is a
> king. Macduff is a noble.
>
> Lady-macbeth persuades Macbeth to want to be king because she is greedy.
> She is able to influence him because he is married to her and because he is
> weak. Macbeth murders Duncan with a knife. Macbeth murders Duncan
> because Macbeth wants to be king and because Macbeth is evil. Lady-macbeth
> kills herself. Macduff is angry. Macduff kills Macbeth because Macbeth
> murdered Duncan and because Macduff is loyal to Duncan.

Next, consider the following exercise:

> Let E be an exercise. E is a story about a weak noble and a greedy lady. The
> lady is married to the noble. In E show that the noble may want to be king.

Told by a teacher that *Macbeth* is to be considered a precedent, it is announced that the
precedent suggests that the noble may want to be king. Then a principle-capturing if-then
rule is created suggesting that the weakness of a noble and the greed of his wife can cause
the noble to want to be king. The rule looks like this:

```
Rule
    RULE-1
if
    [ LADY-4 HQ GREEDY ]
    [ NOBLE-4 HQ WEAK ]
    [ [ NOBLE-4 HQ MARRIED ] TO LADY-4 ]
then
    [ NOBLE-4 WANT [ NOBLE-4 AKO KING ] ]
case
    MA
```

The exercise problem could have been handled by this rule directly, without recourse to the *Macbeth* precedent, were it available when the problem was posed. Thus the rule adds power. Unfortunately, it also adds blunder, as when the following exercise is given:

> Let E be an exercise. E is a story about a weak noble and a greedy lady. The lady is married to the noble. He does not like her. In E show that the noble may want to be king.

This situation is different because we know that it is difficult for a person to influence someone who does not like him. Evidently, the rule is overly general, ready to reach conclusions when it should not.

This paper introduces extensions to the existing theory such that the implemented system behaves correctly on the given example and many others. The improved system works because of the following:

o   The *blocking principle:* Suppose a rule, derived from a precedent, seems to apply to a problem. Consider the relations in that part of the precedents's causal structure involved in forming the rule. If any such relation corresponds to a relation that is either false or manifestly implausible in the problem situation, then the rule based on the precedent does not apply.

o   The *prima facie conjecture:* A relation is manifestly implausible if its negation can be shown by a direct, one-step inference from relations already in place.
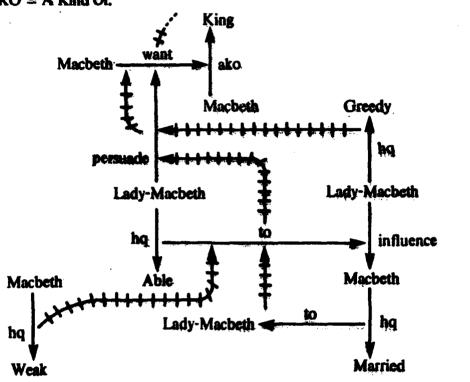
Thus the improved system works, satisfying one important criterion for success, and it works because it exploits identifiable ideas, satisfying another.   ·
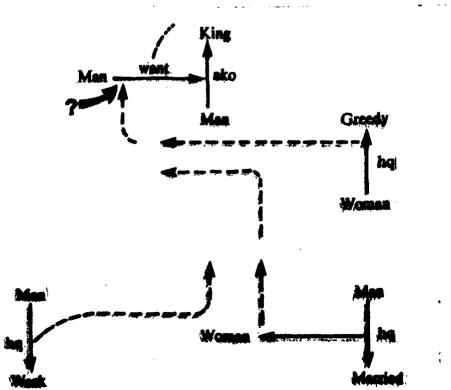
REASONING AND CREATING RULES USING ANALOGY

Let us review how rules are generated. Consider the *Macbeth* precedent, given earlier, together with the exercise, both expressed in semantic-network form, as shown in figure 1. When asked to demonstrate that the man may want to be king, given the *Macbeth* precedent, the system proceeds as follows:

Figure 1: Problems are solved by transferring the existing cause relations of a precedent (crossed lines, part a) onto the problem to be solved (dotted lines, part b). HQ = Has Quality. AKO = A Kind Of.

o   The people in the precedent are matched with the people in the exercise. More
     generally, precedent parts are matched with exercise parts.

o   The causal structure of the precedent is mapped onto the exercise.

o   It is determined that the mapped causal structure ties the relation to be shown to
     relations known to be true.

o   A rule is constructed, with generalizations of the exercise relations used becoming *if*
     parts and a generalization of the relation to be shown becoming the *then* part.

When a single precedent cannot supply the total causal structure needed, the system
attempts to chain several together. In the example, if it were not known already that the
woman is greedy, as required for application of the *Macbeth* precedent, greed might be
established through another precedent or already-learned rule. A previous paper explains
this in detail [Winston 1981].

IMPROVING PERFORMANCE BY ENABLING CENSORS

So far we have established that rules can be generated and that they need to be blocked in
certain circumstances. There are three obvious ways to arrange for blocking:

> First, expand the *if* part of an offending rule, restricting its use. One problem with
this idea is that rules can become bloated with endless tests for increasingly unlikely
minutiae. Such bloat makes rules obscure and hard to criticize, debug, and improve, for
both us people and for reasoning programs.

> Second, attach censors to each rule. Have the censors check the problem to be solved
for contraindications to the rules the censors are attached to. One problem is that the rules
can become bloated with censor names; these censor names would give no explicit insight
into when the rules do not apply.

> Third, have censors watch for particular relations. Forbid any rule or precedent to
work toward establishing a relation that a censor objects to. One problem is that the rules
continue to look silly, containing no hint about when they do not apply.

Censors can Block Augmented If-then Rules

A better, less obvious idea, is this:

o   Augment each rule at the time it is generated with entries that correspond to all
     relations in the causal structure lying between relations that enter the *if* part of the
     rule and the relation that enters the *then* part of the rule. These intermediate entries
     constitute the *if-plausible* part of the rule. According to the blocking principle, if any

entry in the *if-plausible* part of the rule corresponds to something that is manifestly improbable, then the rule does not apply.

Clearly a relation is manifestly improbable if the existing facts indicate that the relation is false. But introspectively, it seems unreasonable to go deeply into reasoning about *if-plausible* entries. Hence the implemented system adheres to the following specialization of the *prima facie* conjecture:

o   If any entry in the *if-plausible* part of a rule corresponds to a relation that can be shown to be false by another rule working directly from relations already in place, then block the rule.

Suppose, for example, that a rule's *if-plausible* part requires that someone be able to influence another. Such a rule will be blocked if the person to be influenced does not like the other. The augmented form of RULE-1 is:

```
Rule
    RULE-1
if
    [ LADY-4 HQ GREEDY ]
    [ NOBLE-4 HQ WEAK ]
    [ [ NOBLE-4 HQ MARRIED ] TO LADY-4 ]
if plausible
    [ LADY-4 PERSUADE [ NOBLE-4 WANT [ NOBLE-4 AKO KING ] ] ]
    [ [ LADY-4 HQ ABLE ] TO [ LADY-4 INFLUENCE NOBLE-4 ] ]
then
    [ NOBLE-4 WANT [ NOBLE-4 AKO KING ] ]
```

The blocking rule is:

```
Rule
    RULE-2
if
    [ [ PERSON-8 LIKE PERSON-7 ] HQ FALSE ]
then
    [ [ [ PERSON-7 HQ ABLE ] TO [ PERSON-7 INFLUENCE PERSON-8 ] ] HQ
      FALSE ]
```

A rule becomes a *censor* when it blocks the application of another rule. Since censors look just like any other rules, censors can be learned, stored, and retrieved in the same ways.

Note that when the illustrated rule is used to block another, it only works if it is known at the time of use that there is dislike. There is no attempt to demonstrate dislike when not already known.

Note that the viability of the *prima facie* conjecture depends on having a rich vocabulary of relations. It would be difficult to demonstrate anything in one step if all relations were reduced to canonical constellations of small-vocabulary primitives. This

opens the question of just how rich the vocabulary should be, a question answered operationally by using freely those relations for which there are common natural-language words.

The viability of the *prima facie* conjecture also depends on having all solid facts available before backward-chaining problem solving begins. This means that all solid facts are either given facts or deduced already by forward chaining from given facts using reliable, potentially relevant rules. Reliability is insured by forward chaining only with rules that reach unassailable conclusions. Relevance cannot be insured, but can be rendered more likely. One way is to use the context mechanism described in an earlier paper [Winston 1981].

Censors can Block Censors

Actually, it is possible to be influenced by someone you dislike if for some reason you trust them in spite of the dislike. Perhaps the real able-to-influence censor should look like this:

```
Rule
    RULE-2
if
    [ [ PERSON-8 LIKE PERSON-7 ] HQ FALSE ]
if plausible
    [ [ PERSON-8 TRUST PERSON-7 ] HQ FALSE ]
then
    [ [ [ PERSON-7 HQ ABLE ] TO [ PERSON-7 INFLUENCE PERSON-8 ] ] HQ
      FALSE ]
```

Such a censor could be blocked by another censor which states that you believe someone if they have the ability to convince you:

```
Rule
    CENSOR-1
if
    [ [ PERSON-6 HQ ABLE ] TO [ PERSON-6 CONVINCE PERSON-6 ] ]
then
    [ PERSON-6 TRUST PERSON-6 ]
```

To illustrate how these can interact, consider the following situation:

> Let E be an exercise. E is a story about a weak noble and a greedy lady. The
> lady is married to the noble. He does not like her. The lady is able to convince
> the noble. In E show that the noble may want to be king.

This produces the following scenario:

o   First, the problem is posed and RULE-1 is fetched. Its *if* parts are satisfied.

o   Next, the *if-plausible* part of RULE-1 is examined. The line involving ability to influence causes RULE-2 to be fetched. Its *if* parts are satisfied. RULE-1 is about to be blocked.

o   But RULE-2's *if-plausible* part must be examined. The line involving not believing causes CENSOR-1 to be fetched. Its *if* parts are satisfied. Thus CENSOR-1 blocks RULE-2, preventing RULE-2 from blocking RULE-1.

o   Finally, RULE-1 succeeds, establishing the relation originally asked about.

Experiment 1, in the appendix, gives the actual trace of the system for this example.

Augmented If-then Rules are not Rules of Inference

It is tempting to write censors in the following way:

$$A_1 \wedge ... \wedge A_n \wedge \neg(\neg B_1 \vee ... \vee \neg B_n) \Rightarrow C,$$

or alternatively,

$$A_1 \wedge ... \wedge A_n \wedge B_1 \wedge ... \wedge B_n \Rightarrow C,$$

where the *A*s are in the *if* part of the rule and the *B*s are in the *if-plausible* part.

Logical notation is deceptive, however, for in the use of augmented if-then rules, the *A*s and *B*s get treated differently from each other, in contrast to the conventions of traditional logic: unlimited effort is to be put into showing the *A*s are true; only one-step effort is put into showing that the negation of the *B*s is true, with the *B*s assumed true on failure.

Note that rules used as censors are not permitted to create new objects. This insures that the amount of computation added by the application of censors to *if-plausible* entries is bounded even though censors have their own *if-plausible* parts that must be checked by censors. I believe it is likely that censor computations will prove in practice to be broad and shallow, as well as bounded, suggesting parallel implementation.

Augmented Rules suggest an Approach to certain Definition Problems

Winograd has discussed the difficulty of definition using the word *bachelor* [Winograd 1976]. To be sure, a bachelor is an unmarried adult man, but Winograd notes that such a definition can cause trouble if used when someone says, "Please invite some

achelors to my party," for it would be strange to invite certain kinds of bachelors.
tample, Catholic priests and misogamists, while satisfying the dictionary definition,
:arly not what a party giver has in mind.

Since the exception possibilities seem limitless, Winograd feels it is inappropriate to
definition of *bachelor* on a clearly defined, small set of primitive propositions,
ig that it is better to think of using some abstract measure of closeness to an extensible
ſ exemplars.   Woods takes issue with Winograd's view, feeling that correct
standing must involve an explicit selection of a particular word sense, rather than
ess to a generally applicable exemplar set [Woods 1981].

The augmented-rule idea may offer a slightly different approach to the problem.
der the following definition of *bachelor*, stated as an augmented if-then rule:

UI.E-2

```
MAN-10 AKO MAN ]
[ MAN-10 HQ MARRIED ] HQ FALSE ]
MAN-10 AKO ADULT ]
ausible
[ MAN-10 HQ MARRIED ] HQ EXPECTED ]
[ MAN-10 HQ ABLE ] TO [ MAN-10 HQ MARRIED ] ]

MAN-10 AKO BACHELOR ]
```

this definition, the conclusion can be avoided, even though the *if* part of the rule is
:atisfied, providing that the individual involved is not able to be married or is not
ted to be married. This takes care of the priest and the misogamist problems, given
llowing censors:

ENSOR-1

```
PERSON-1 AKO MISOGAMIST ]

[ [ PERSON-1 HQ MARRIED ] HQ EXPECTED ] HQ FALSE ]
```

ENSOR-2

```
MAN-4 AKO PRIEST ]

[ [ MAN-4 HQ ABLE ] TO [ MAN-4 HQ MARRIED ] ] HQ FALSE ]
```

itly, it is possible to have a simple, stable definition of bachelor, while at the same
llowing for knowledge relevant to bachelors to interact with the definition, when
priate, as that knowledge is accumulated. As more is learned, the definition is used
ntelligently, and, in a sense, the definition is never closed.

How does capturing the meaning of *bachelor* with an augmented if-then rule compare with other approaches? One point of view is that Winograd's exemplars correspond to rule-generating precedents, and learned augmented if-then rules correspond to Woods's selectable word senses. We will turn to learning about bachelors from precedents in a moment.

Censors can Improve Precedent Reasoning

While censors were originally investigated in this work in order to cure the apparent silliness of some learned rules, they help in another context too. When ordinary precedent-exercise problem solving is in progress, the analogy part of the system works back through the causal structure in the precedent, looking for relations that correspond to relations in the exercise. Each time there is no corresponding relation, before the system moves further through the causal structure, it does a censor check.

o   If a relation is encountered in the causal structure of the precedent that corresponds to a relation that is manifestly improbable in the exercise, then the precedent cannot support a conclusion.

LEARNING AUGMENTED RULES

Since censor rules and definition rules are just rules used in a special way, they can be learned just like any other rules. This may be by direct telling, or it may be by precedent and exercise, or it may be by near-miss.

Augmented Rules can be Learned by Precedent and Exercise

Here is a precedent and an exercise for learning the bachelor definition rule:

> Let S be a story. S is a story about Casanova. Casanova is a bachelor because he is a man and because he is expected to be married. He is expected to be married because he is able to be married. He is able to be married because he is an adult and because he is not married.

> Let E be an exercise. E is a story about Henry. He is a man and an adult. He is not married. In E show that Henry is a bachelor.

A trace of the learning system acting on this precedent exercise pair is given in Experiment 2 of the appendix.

Of course, one might argue that providing the precedent involving Casanova is unrealistic spoon feeding. Indeed, it may well be, so it is important to understand that the same bachelor rule can be learned using several independent precedents:

Let S be a story. S is a story about a man. He is a bachelor because he is expected to be married. He is a bachelor because he is a man.

Let S be a story. S is a story about a man. He is expected to be married because he is able to be married.

Let S be a story. S is a story about a man. He is able to be married because he is an adult and because he is not married.

Alternatively, the bachelor rule can be learned using several previously-learned rules:

```
Rule
    STORY-1
if
    [ MAN-1 AKO MAN ]
    [ [ MAN-1 HQ MARRIED ] HQ EXPECTED ]
then
    [ MAN-1 AKO BACHELOR ]

Rule
    STORY-2
if
    [ [ MAN-2 HQ ABLE ] TO [ MAN-2 HQ MARRIED ] ]
then
    [ [ MAN-2 HQ MARRIED ] HQ EXPECTED ]

Rule
    STORY-3
if
    [ MAN-3 AKO ADULT ]
    [ [ MAN-3 HQ MARRIED ] HQ FALSE ]
then
    [ [ MAN-3 HQ ABLE ] TO [ MAN-3 HQ MARRIED ] ]
```

Also, it is possible to learn a rule that allows a married Moslem, seeking an additional wife, to be considered a bachelor. See Experiment 3 in the appendix.

Augmented Rules can be Learned by Near-miss

Of course, there should be some way of recovering if an impoverished definition is acquired early on. The near-miss idea seems useful in such situations. Consider this scenario:

o   A teacher tells the system that a bachelor is an unmarried, adult man. This produces an impoverished definition of bachelor, one without anything in the *if-plausible* part.

o   The teacher complains when the system identifies a Catholic priest as a bachelor.

o   The system notices that the only robust difference between the priest and other people who are correctly identified as bachelors is that the priest is not able to be married.

o   The system guesses that bachelors must be able to be married and puts an appropriate entry in the *if-plausible* part of the bachelor definition.

Experiment 4 in the appendix shows the system going through such a scenario, producing an improved definition of bachelor. Of course, this is a particularly simple situation since there is but one object involved and the descriptions are such that the near-miss-causing relation is the only relation that is caused by something and not deemed plausible in a situation where the rule does apply. It is not known how difficult it would be to identify the right difference in general. Recent work by Berwick on syntax acquisition [1982] and by Minsky in concept learning [unpublished draft] suggest that if it is difficult to identify the right difference, a learning system should simply give up, waiting for more transparent examples to come along.

THE IMPLEMENTED SYSTEM

The example precedents, exercises, rules, and censors in this paper are shown in the exact English form used by the implemented system. Translation from English into the semantic net representation used by the system is done by a parser developed and implemented by Boris Katz [Katz 1980, Katz and Winston 1982]. The grammar used by the parser is also used by a generator, which produces English versions of the rules. For example, the generator converts

```
Rule
    RULE-2
if
    [ MAN-10 AKO MAN ]
    [ [ MAN-10 HQ MARRIED ] HQ FALSE ]
    [ MAN-10 AKO ADULT ]
if plausible
    [ [ MAN-10 HQ MARRIED ] HQ EXPECTED ]
    [ [ MAN-10 HQ ABLE ] TO [ MAN-10 HQ MARRIED ] ]
then
    [ MAN-10 AKO BACHELOR ]

into
```

Rule-2 concerns a man. If the man is not married and he is an adult, then he is a bachelor, assuming he is expected to be married and he is able to be married.

Representative traces of the system in action, along with the English-form of the data used, are given in the appendix. So far, the system knows a few dozen censors, most of which it is told, all of which it can learn from precedents or rules and exercises. Clearly the number is enough to do surface-scratching experiments and to illustrate the ideas, but an order of magnitude or two more will be required to demonstrate the ideas.

OPEN QUESTIONS

It is plain that this work is only a beginning. Work is in progress on several related fronts:

o   In collaboration with Boris Katz: the problem of retrieving precedents from a data base so that they need not be given by a teacher.

o   In collaboration with Tomas O. Binford (Stanford University), Michael Lowry (Stanford University), and Boris Katz: the problem of creating appearance descriptions from functional descriptions, precedents, and examples.

o   In response to a suggestion by J. Michael Brady: an augmentation of the rules with an *if-relevant* part in addition to the *if-plausible* part described in this paper. The idea is that the *if-relevant* part will somehow keep track of the ultimate goals a rule may be relevant to, so that the rule is used in forward chaining only if one of the potential ultimate goals is involved in the problem to be solved. This would make the rules look like this in logical notation:

$$A_1 \wedge ... \wedge A_n \wedge \neg(\neg B_1 \vee ... \vee \neg B_n) \wedge (G_1 \vee ... \vee G_n) \Rightarrow C,$$

where the *A*s are in the *if* part of the rule, the *B*s are in the *if-plausible* part, and the *G*s are in the *if-relevant* part; and where it is understood that only one-step effort is to be put into the *B*s and *G*s.

This would complement the existing context mechanism explained previously [Winston 1981].

In addition, the following open questions, enumerated in a previous paper, remain open [Winston 1981]:

o   There is no way to handle degree of certainty of cause. Moreover, there is no way to handle subcategories of cause such as those sketched by Rieger [1978].

o   There is no way to handle constraints about quantities such as those constraints that appear in the work of Forbus [1982].

o   There is no way to summarize an episode in a story so as to make a general precis leading to more abstract rules. Lehnert's summarization work should be tried [Lehnert 1981].

o   There are no satisfying ideas about the role of abstraction in doing matching and indexing and retrieving.

o   The representation for time is impoverished. Similarly quantification, negation, disjunction, and perspective are missing.

CONCLUSION: SIMPLE IDEAS HAVE PROMISE

This paper is about a set of ideas that enable improvement in the reliability of learned rules. The extended theory enables improved performance in those domains subject to problem solving by analogy. Such domains satisfy certain restrictions:

o   The situations in the domain can be represented by the relations between the parts together with the classes and properties of those parts.

o   The importance of a part of a description is determined by the constraints it participates in.

o   Constraints that once determine something will tend to do so again.

Things that involve spatial, visual, and aural reasoning do not seem to satisfy all the restrictions. Things that involve Management, Political Science, Economics, Law, Medicine, and ordinary common sense do seem to satisfy the restrictions, however, and are targets for the learning and reasoning ideas of the theory:

o   Actor-object representation.

o   Importance-dominated matching.

o   Analogy-based reasoning using constraint transfer.

o   If-then rules learned by solving problems.

o   If-then rules improved by modifications based on near misses.

o   If-then rules augmented by if-plausible parts.

o   Blocking censors that create fences around rules using *prima facie* evidence.

RELATED WORK

This work builds on the MACBETH system [Winston 1980, 1982], which concentrated on analogy and rule acquisition. Also, Minsky's views on censors had a major influence [Minsky 1980]. To a lesser extent, the idea of learning by near miss is involved [Winston 1970].

The augmented if-then rule is a special case of the annotated if-then rule introduced by Goldstein and Grimson in a paper on flight simulation [1977]. They had the idea that if-then rules should have if-plausible conditions (which they called caveats), as well as rationales, plans, and control information. The work of Brown and VanLehn on explaining subtraction bugs is a more recent precedent for using censors to block rules, although their censors (which they call critics) are triggered by what a rule does, rather than by if-plausible conditions [Brown and VanLehn 1980].

The idea that censors should work only with the facts in hand is a variant on the theme of reasoning using limited resources, an idea that is discussed widely, particularly in the expert-systems literature.

John Mallery observed in conversation that the definition of *bachelor* really should say something about being expected to be married, stimulating me to try handling the bachelor problem within the *if-plausible* framework. Boris Katz pointed out that the *prima facie* conjecture does not make sense unless all reliable, potentially relevant forward chaining is done first.

ACKNOWLEDGMENTS

REFERENCES

Berwick, Robert "Locality Principles and the Acquisition of Syntactic Knowledge," PhD Thesis, Department of Electrical Engineering & Computer Science, MIT, 1982.

Brown, John Seely and Kurt VanLehn, "Repair Theory: A Generative Theory of Bugs in Procedural Skills," *Cognitive Science*, vol. 4, no. 4, October-December, 1980.

Davis, Randall "Applications of Meta Level Knowledge to the Construction, Maintenance, and Use of Large Knowledge Bases," Ph. D. Thesis, Stanford University, Stanford, California, 1979. Also in *Knowledge-Based Systems in Artificial Intelligence*, Randall Davis and Douglas Lenat, 1980.

Forbus, Ken, "Qualitative Reasoning about Physical Processes," submitted for publication, 1982.

Goldstein, Ira P. and Eric Grimson, "Annotated Production Systems A Model for Skill Acquisition," *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, 1977. Available through Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania, 15213.

Katz, Boris and Patrick H. Winston, "Parsing and Generating English using Commutative Transformations," Artificial Intelligence Laboratory Memo No. 677, May, 1982. Also see "A Three-Step Procedure for Language Generation," by Boris Katz, Artificial Intelligence Laboratory Memo No. 599, December, 1980.

Lehnert, Wendy, "Plot Units and Narrative Summarization," *Cognitive Science*, vol 4, 1981.

Minsky, Marvin, "Jokes and the Logic of the Cognitive Unconscious," M.I.T. Artificial Intelligence Laboratory Memo No. 603, November 1980.

Minsky, Marvin, unpublished draft on concept learning.

Rieger, Chuck, "On Organization of Knowledge for Problem Solving and Language Comprehension," *Artificial Intelligence*, vol. 7, no. 2, 89-128, 1978.

Winograd, Terry, "Towards a Procedural Understanding of Semantics," Stanford Artificial Intelligence Laboratory Memo AIM-292, Stanford Computer Science Department Report No. STAN-CS-76-580, November, 1976.

Winston, Patrick Henry, "Learning Structural Descriptions from Examples," Ph.D. thesis, MIT, 1970. A shortened version is in *The Psychology of Computer Vision*, edited by Patrick Henry Winston, McGraw-Hill Book Company, New York, 1975.

Winston, Patrick Henry, "Learning by Creating and Justifying Transfer Frames," *Artificial Intelligence*, vol. 10, no. 2, 147-172, 1978.

Winston, Patrick Henry, "Learning and Reasoning by Analogy," *CACM*, vol. 23, no. 12, December, 1980. A version with details is available as "Learning and Reasoning by Analogy: the Details," M.I.T. Artificial Intelligence Laboratory Memo No. 520, April 1979.

Winston, Patrick Henry, "Learning New Principles from Precedents and Exercises," to appear in *Artificial Intelligence*. A version with details is available as "Learning New Principles from Precedents and Exercises: the Details," M.I.T. Artificial Intelligence Laboratory Memo No. 632, May 1981.

Woods, William A., "Procedural Semantics as a Theory of Meaning," Bolt Beranek and Newman Report No. 4627, March 1981. Also available in *Computational Aspects of Linguistic Structures*, A. Joshi, I. Sag, and B. Webber (eds.), Cambridge University Press.

APPENDIX: THE EXPERIMENTS

This appendix contains traces of the set of experiments which are described in the body of this paper.

Experiment 1:

This exercise illustrates the use of one censor to block another, preventing the blocking of a rule.

> Let E be an exercise. E is a story about a weak noble and a greedy lady. The lady is married to the noble. He does not like her. The lady is able to convince the noble. In E show that the noble may want to be king.

```
I am trying to show [ NOBLE-3 WANT [ NOBLE-3 AKO KING ] ]
Supply y, n, ?, r = rules, p = precedents, or a suggestion:
> r
I find RULE-1 +
Matching EXERCISE-4 to RULE-1 producing MATCH-22
0+
The match score is 100. %
((LADY-3 LADY-4) (NOBLE-3 NOBLE-4))
I note that [ LADY-3 HQ GREEDY ] for use with RULE-1
I note that [ NOBLE-3 HQ WEAK ] for use with RULE-1
I note that [ [ NOBLE-3 HQ MARRIED ] TO LADY-3 ] for use with RULE-1
----- start of censor check -----
++
Matching EXERCISE-4 to RULE-2 producing MATCH-23

The match score is 100. %
((NOBLE-3 PERSON-8) (LADY-3 PERSON-7))
I note that [ [ NOBLE-3 LIKE LADY-3 ] HQ FALSE ] for use with RULE-2
----- start of censor check -----
++
Matching EXERCISE-4 to CENSOR-1 producing MATCH-24

The match score is 100. %
((LADY-3 PERSON-6) (NOBLE-3 PERSON-5))
I note that [ [ LADY-3 HQ ABLE ] TO [ LADY-3 CONVINCE NOBLE-3 ] ]
  for use with CENSOR-1
The evidence from CENSOR-1 prevents [ NOBLE-3 TRUST LADY-3 ]
----- end of censor check -----
Rule RULE-2 is blocked.
The evidence from RULE-2 does not prevent [ [ LADY-3 HQ ABLE ] TO [
  LADY-3 INFLUENCE NOBLE-3 ] ]
----- end of censor check -----
The evidence from RULE-1 indicates [ NOBLE-3 WANT [ NOBLE-3 AKO KING ]
  ]
No need to derive a new rule.
```

Experiment 2:

The task is to learn what a bachelor is.

> Let S be a story. S is a story about Casanova. Casanova is a bachelor because he is expected to be married. Casanova is a bachelor because he is a man. He is expected to be married because he is able to be married. He is able to be married because he is an adult and because he is not married.

> Let E be an exercise. E is a story about Henry. He is a man and an adult. He is not married.

> In E show that Henry is a bachelor.

```
I am trying to show [ HENRY AKO BACHELOR ]
Supply y, n, ?, r = rules, p = precedents, or a suggestion:
> story-1
Matching EXERCISE-1 to STORY-1 producing MATCH-9
0+
The match score is 100. %
((HENRY CASANOVA))
----- start of censor check -----
+
Matching EXERCISE-1 to CENSOR-1 producing MATCH-10

The match score is 100. %
((HENRY PERSON-1))
The evidence from CENSOR-1 does not prevent [ [ HENRY HQ MARRIED ]
  HQ EXPECTED ]
----- end of censor check -----
----- start of censor check -----
+
Matching EXERCISE-1 to CENSOR-2 producing MATCH-11

The match score is 100. %
((HENRY MAN-4)) .
The evidence from CENSOR-2 does not prevent [ [ HENRY HQ ABLE ] TO [
  HENRY HQ MARRIED ] ]
----- end of censor check -----
I note that [ HENRY AKO ADULT ] for use with STORY-1
I note that [ [ HENRY HQ MARRIED ] HQ FALSE ] for use with STORY-1
I note that [ HENRY AKO MAN ] for use with STORY-1
The evidence from STORY-1 indicates [ HENRY AKO BACHELOR ]
Rule RULE-2 is derived from STORY-1 and looks like this:

Rule
    RULE-2
if
    [ MAN-10 AKO MAN ]
```

```
     [ [ MAN-10 HQ MARRIED ] HQ FALSE ]
     [ MAN-10 AKO ADULT ]
if plausible
     [ [ MAN-10 HQ MARRIED ] HQ EXPECTED ]
     [ [ MAN-10 HQ ABLE ] TO [ MAN-10 HQ MARRIED ] ]
then
     [ MAN-10 AKO BACHELOR ]
case
     STORY-1

Should I index it as a rule?
> y
```

Experiment 3:

The task is to learn an alternative definition of bachelor, suited to those who may have more than one wife.

> Let S be a story. S is a story about Casanova. Casanova is a bachelor because he is expected to be married. Casanova is a bachelor because he is a man. He is expected to be married because he is able to be married. He is able to be married because he is an adult and because he is not married.

> Let E be an exercise. E is a story about a man. He is an adult and a moslem. He is married. He is able to be married again because he is a moslem.

> In E show that the man is a bachelor.

```
I am trying to show [ MAN-7 AKO BACHELOR ]
Supply y, n, ?, r = rules, p = precedents, or a suggestion:
> story-1
Matching EXERCISE-4 to STORY-1 producing MATCH-12
+
The match score is 100. %
((MAN-7 CASANOVA))
----- start of censor check -----
+
Matching EXERCISE-4 to CENSOR-1 producing MATCH-13

The match score is 100. %
((MAN-7 PERSON-1))
The evidence from CENSOR-1 does not prevent [ [ MAN-7 HQ MARRIED ] HQ
   EXPECTED ]
----- end of censor check -----
I note that [ [ MAN-7 HQ ABLE ] TO [ MAN-7 HQ MARRIED ] ] for use with
   STORY-1
I note that [ MAN-7 AKO MAN ] for use with STORY-1
The evidence from STORY-1 indicates [ MAN-7 AKO BACHELOR ]
```

```
Rule RULE-3 is derived from STORY-1 and looks like this:

Rule
    RULE-3
if
    [ MAN-11 AKO MAN ]
    [ [ MAN-11 HQ ABLE ] TO [ MAN-11 HQ MARRIED ] ]
if plausible
    [ [ MAN-11 HQ MARRIED ] HQ EXPECTED ]
then
    [ MAN-11 AKO BACHELOR ]
case
    STORY-1

Should I index it as a rule?
>  y :
```

Experiment 4:

The task is to improve a simple definition of bachelor using a near miss.

Let R be a rule. R is a story about a person. The person is a bachelor because he is a man, because he is an adult, and because he is not married.

Let X be a story. X is a story about a man. He is an adult. He is not married.

Let Y be a story. Y is a story about a man. He is an adult. He is not married. He is not able to be married because he is a priest.

Fix rule R which applies to X but not to Y.

```
The raw rule is:
Rule
    RULE-1
if
    [ PERSON-2 AKO MAN ]
    [ PERSON-2 AKO ADULT ]
    [ [ PERSON-2 HQ MARRIED ] HQ FALSE ]
then
    [ PERSON-2 AKO BACHELOR ]
Matching STORY-6 to STORY-5 producing MATCH-14

The match score is 0. %
NIL
Match for MAN-9 using AKO is MAN-8
----- start of censor check -----
+
Matching STORY-6 to CENSOR-2 producing MATCH-16
```

```
The match score is 100. %
((MAN-8 MAN-4))
The evidence from CENSOR-2 does not prevent [ [ MAN-8 HQ ABLE ] TO [
  MAN-8 HQ MARRIED ] ]
----- end of censor check -----
Is it plausible that [ [ [ MAN-8 HQ ABLE ] TO [ MAN-8 HQ MARRIED ] ]
  HQ FALSE ]
> n
Evidently the relevant difference is that [ [ [ MAN-9 HQ ABLE ] TO [
  MAN-9 HQ MARRIED ] ] HQ FALSE ] in STORY-6
Matching STORY-6 to RULE-1 producing MATCH-18
+
The match score is 100. %
((MAN-9 PERSON-2))
Rule
    RULE-1
if
    [ PERSON-2 AKO MAN ]
    [ PERSON-2 AKO ADULT ]
    [ [ PERSON-2 HQ MARRIED ] HQ FALSE ]
if plausible
    [ [ PERSON-2 HQ ABLE ] TO [ PERSON-2 HQ MARRIED ] ]
then
    [ PERSON-2 AKO BACHELOR ]
```

E

ED

82